Manuscript of:
Szczypiński, P. M., Klepaczko, A., & Zapotoczny, P. (2015)
Identifying barley varieties by computer vision.
Computers and Electronics in Agriculture, 110, 1-8.

http://dx.doi.org/10.1016/j.compag.2014.09.016

Identifying barley varieties by computer vision

Piotr M. Szczypiński [a, *], Artur Klepaczko [a], Piotr Zapotoczny [b]

[a] Institute of Electronics, Lodz University of Technology,
Wolczanska 211/215, 90-924 Lodz, Poland

[b] Department of Agri-Food Process Engineering, University of Warmia and Mazury in Olsztyn,
Heweliusza 14, 10-718 Olsztyn, Poland

Keywords: Barley classification, Computer vision, Machine learning

Abstract

Visual discrimination between barley varieties is difficult, and it requires training and experience. The development of automatic methods based on computer vision could have positive implications for the food processing industry. In the brewing industry, varietal uniformity is crucial for the production of high quality malt. The varietal purity of thousands of tons of grain has to be inspected upon purchase in the malt house.

This paper evaluates the effectiveness of identification of barley varieties based on image-derived shape, color and texture attributes of individual kernels. Varieties can be determined by means of discriminant analysis, including reduction of feature space dimensionality, linear classifier ensembles and artificial neural networks, with high balanced accuracy ranging from 67% to 86%. The study demonstrated that classification results can be significantly improved by standardizing individual kernel images in terms of their anteroposterior and dorsoventral orientation and performing additional analyses of wrinkled regions.

1. Introduction

Barley is a major cereal grain used for both human consumption and animal feed. A total of 47 spring barley varieties (30 malting varieties and 17 fodder varieties) and 17 winter barley varieties are currently included in the Polish National List. Since certain varieties are suitable for specific applications, adequate selection is a crucial step in barley grain processing. For instance, barley grain used for malting should be characterized by low protein content, uniform size and high quality. Therefore, grain has to be controlled for varietal purity and technological quality at every stage of processing (Hulasare et al., 2003).

There are several methods of cereal grain testing, including immunological analysis, DNA analysis, high-performance liquid chromatography, protein electrophoresis and isoenzyme analysis.

* Corresponding author: piotr.szczypinski@p.lodz.pl, tel.: +48 426312642; fax: +48 426362238

Most of those methods are labor intensive and expensive, and the analyses can only be performed by specialized laboratories. An alternative approach involves visual evaluation of grain for varietal classification in line with the International Rules for Seed Testing developed by the International Seed Testing Association (ISTA). The following physical parameters are generally evaluated during a visual inspection: kernel color, kernel shape, shape of lemma base, ventral crease hairs, rachis hairs, teeth on lateral dorsal nerves, wrinkling of the lemma and palea, shape and hairs of lodicules. This technique is much easier to apply than chemical methods, but the reliability of visual evaluation is largely dependent on the skills and experience of the evaluator.

Computer image analysis, a rapid and low-cost technique, is an alternative method that evaluates selected physical attributes of kernels. Varietal classification of cereal grains by computer vision has been widely discussed in literature (Mebatsion et al., 2013; Neuman et al., 1987; Visen et al., 2002; Zapotoczny, 2011a, 2011b). However, none of the cited studies involved correction of image orientation, and classification models relied on analyses of whole kernel regions only. By contrast, an algorithm (Szczypiński and Zapotoczny, 2012) was developed to analyze kernel images in terms of the dorsoventral (the side with or without the crease is visible) and anteroposterior (germ-brush direction along the semi-symmetric axis) orientation of kernels. Kernel images were segmented into specific areas that were inspected individually. The proposed procedure was motivated by the assumption that the shape and properties of a kernel's wrinkled region are genetically determined. A kernel's ventral and dorsal surfaces could have different properties that should be analyzed by distinct algorithms.

Image classification usually involves computation of image attributes (features). An image attribute is a numeric quantity that characterizes the image or its fragment. Numerous algorithms for feature computation focus on different aspects of image appearance, such as brightness and color distribution, texture, shape and topology of a region. Those attributes can be examined in cereal analysis. Neuman et al. (1989) analyzed the color attributes of cultivars belonging to different wheat classes and demonstrated significant differences between varieties. Paliwal et al. (1999) relied on color attributes and shapes described by the Fourier transform of a radial function to distinguish between kernels of wheat, barley, oats and rye. Image histogram parameters computed from monochrome images of a wheat bulk sample proved to be useful for the estimation of moisture content (Manickavasagan et al., 2008). Zielinska et al. (2012) used image texture attributes and morphological (shape) properties to classify four red clover cultivars.

Thousands of attributes can be used to characterize an individual image, depending on its specific application (Szczypinski et al., 2009). Thus, every analyzed image is described by attribute vectors with thousands of dimensions. A statistical analysis is then carried out to examine the distribution of attribute vectors in high-dimensional spaces that are not easy to handle. For instance, overfitting may take place in highly complex models where the number of parameters significantly exceeds the number of observations (Berthold and Hand, 2003). This leads to poor predictive performance as the classifier trained on a set of examples fails to correctly recognize other data. In cereal classification, the artificial neural network is often a method of choice (Goyal, 2013). Neural network classifiers such as multilayer perceptrons allow for nonlinear decision boundaries, and they easily fit exemplary data if the hidden layer contains a large number of neurons. Unfortunately, a higher number of neurons increases the risk of overtraining (Jain et al., 2000), which also leads to poor predictive performance.

The goal of this paper was to evaluate the feasibility of image-driven classification of 11 barley varieties and to identify subsets of attributes with the highest discriminatory ability. The study also set out to establish whether information related to a kernel's wrinkled subregion and its dorsoventral and germ-brush orientation improves classification performance.

2. Materials and Methods

Grain samples used in this study were obtained from selected farms in the Region of Warmia and Mazury, NE Poland. The experimental material consisted of 11 varieties of two-rowed barley: Afrodita, Blask, Bordo, Conchita, Kormoran, Mercanda, Prymus, Serwal, Signora, STH and Victoriana. The differences between kernels belonging to the evaluated varieties are unlikely to be identified by an untrained assessor. Effective discrimination between varieties requires training and experience.

In accordance with ISTA recommendations, grain should be rinsed with water before a sensory evaluation to expose its texture and color. In this study, grain was not rinsed to improve its surface condition because in line with the applied methodology, unwashed grain is more difficult to classify. Therefore, the proposed methodology would be validated more effectively if tested on unwashed grain.

The images were acquired in color at the resolution of 400 dpi, 24 bits per pixel, in the Epson 4990 flatbed scanner. The scanner was placed in a compartment covered with black velvet and its top cover was removed. The kernels were positioned manually using a template with holes punctured specifically for this purpose. When all holes were filled with kernels, excess kernels and, consequently, the template were removed. This technique supported uniform spacing of 450 kernels per scan. However, individual kernels can be rotated, and they can produce images of the dorsal or the ventral side. The resulting images show relatively bright and disjoint kernel areas on a dark background (Fig. 1). In total, 33 scanned images of 11 barley varieties were acquired, with three images per variety, which resulted in more than 13000 individual objects for analysis.

One of the objectives was to determine whether identification of dorsoventral orientation, correction of the germ-brush direction and supplementary analysis of wrinkled subregions would improve classification results. The other goal was to identify attributes that best discriminate barley varieties. The experiment had the following design: kernels were identified in images and described individually in terms of their numerical properties – attributes. Images of individual kernels were rotated with respect to the germ-brush direction and split into two groups based on their dorsoventral orientation. The resulting images were repeatedly numerically characterized. Wrinkled kernel areas were determined and their attributes were computed to supplement the characteristics of rotated kernel images. The analysis produced data sets characterizing kernel images at three increasingly advanced preprocessing levels. A discriminant analysis was performed separately on the three data sets, and it involved attribute selection and classification. The analysis revealed kernel properties that were most suitable for varietal discrimination and demonstrated whether the proposed image processing method affected classification performance.

Analytical steps are explained in detail in the following subsections, including image processing to identify individual kernels and determine their orientation (section 2.1), computation of image features to numerically characterize every object's texture, color and shape (section 2.2), and supervised learning to select attributes that best discriminate barley varieties and contribute to high performance classification (section 2.3).

2.1. Image preprocessing

An image is processed to increase its suitability for the following analytical step, which is the computation of kernel attributes. Morphological features strictly depend on the region shape. Brightness, color and texture differ between kernel regions and the image background. For these reasons it is crucial to accurately outline image regions corresponding to the kernels, to correctly estimate the attributes. Selected attributes are rotation dependent – their values change as the image is rotated. The above poses problems in descriptions of kernel geometry or texture, since rotation

dependent attributes may be incorrectly estimated for unaligned, diversely rotated objects. Attributes characterizing brightness distribution also differ between kernels viewed from the dorsal or the ventral side. Those differences may hamper varietal classification.

To address the above issues, we have designed an image preprocessing algorithm that involves segmentation, identification of kernel region images, determination of kernel orientation and adjustment of germ-brush alignment (Fig. 1). The algorithm has been described in detail by Szczypiński and Zapotoczny (2012). The essentials are briefly recalled in this paper.

The first step is image segmentation, which splits image pixels into two subsets: image background and kernel areas. Since kernel regions are brighter than the background, the image gray-level thresholding method is used – pixels darker than the experimentally established threshold are assigned to the first subset and the remaining ones are allocated to the second subset. Image elements brighter than the threshold form disjoint regions that correspond to individual kernels. Such regions are not always uniformly connected, they may contain undesirable holes, and their contours are usually rough. To address these problems, morphological opening and closing algorithms are applied. The opening operator removes small-sized, peninsula-shaped remnants located near contours, whereas the closing algorithm removes gulf-shaped cavities and holes. The isolated regions corresponding to individual kernels are then identified and stored. The preconditioned images constitute the *S* data set.

Kernel orientation is determined in the next step. Two issues are addressed at this point. Firstly, the angle of the anteroposterior axis is determined, and the image is rotated by that angle. Secondly, a kernel's visibility is determined from dorsal and ventral sides, and the images are split into two subsets accordingly. To determine the germ-brush direction, the boundary of every kernel is approximated by an ellipse whose longer diameter roughly designates the anteroposterior axis. Barley kernels are wider on the germ side and narrower on the other side. Therefore, germ and brush sides are determined by measuring kernel width on both sides of an ellipse's shorter diameter. Finally, the dorsoventral orientation is established by crease detection. The crease forms a dark elongated area across the mid-line of the grain that can be distinguished by grey-level thresholding. All images from the *S* data set are rotated in the germ-brush direction and split into two subsets: *V* – showing the ventral side, and *D* – showing the dorsal side of a kernel.

The last stage of image preprocessing involves the identification of wrinkled regions. The regions are discriminated based on their fine-grained texture, a mixture of relatively small dark and bright patches, whereas the remaining part of the kernel surface looks relatively smooth (Fig. 2). A gray-level run-length analysis was carried out to identify wrinkled regions (Haralick, 1979). The concept is based on the observation that vertical pixel runs of similar gray-levels are shorter in wrinkled areas, whereas longer pixel runs are observed in smooth areas. The run length attribute is mapped onto the image, where low attribute values indicate the wrinkled area and higher values denote the smooth area. Kernel images from *V* and *D* sets supplemented with information about the corresponding wrinkled regions constitute two data sets denoted as *V*+ and *D*+, respectively.

2.2. Computation of attributes

An image attribute, or a feature, is a numerical indicator that describes the properties of an image or its region. In the attribute computation procedure, the image is analyzed to determine the value of a numerical indicator. Various algorithms have been proposed for computing attributes that characterized different aspects of a region, including shape, brightness, texture or color. To describe an image appearance comprehensively, a feature vector (a certain number of ordered attributes) is used. In successive stages of image analysis, such vector distribution may be examined by pattern recognition methods.

A total of 54 morphological features were used in this study to describe the shape of an image region (Majumdar and Jayas, 2000; Pavlidis, 1980; Zapotoczny, 2011b). They included estimations

of a region's area, height, width, perimeter, maximum and minimum diameters, etc. The results of those measurements are used to calculate more complex factors, such as aspect ratio, area ratio, circularity, elongation, slenderness, compactness, corrugation, Malinowska's ratio, Danielson index and Blair-Bliss ratio. Morphological features also include geometrical moments, including the center of gravity, as well as metrics describing the distance between an object's perimeter and its center. The width, height and area of inscribed and circumscribed squares, circles and ellipses are also estimated.

Texture can be determined in terms of spatially arranged complex pixel patterns that are somewhat homogeneous in appearance. Texture attributes can be determined with the use of three computation approaches: statistical, model-based and image-transform. In the statistical approach, texture is represented by non-deterministic properties of gray-level distributions and spatial relationships. Statistical metrics include the mean, variance, skewness or percentiles of the original image histogram. The attributes derived from the grey-level co-occurrence matrix (Haralick et al., 1973) and the grey-level run-length matrix (Haralick, 1979) are also computed. In the model-based approach, texture is interpreted with the use of generative image models or stochastic models (Jain, 1989; Materka and Strzelecki, 1998). The method used in this study relies on the parameters of an auto-regressive model, where pixel intensity is predicted as the weighted sum of four neighboring pixel intensities. The modeled parameters are the weights and the variance of a prediction error. Transform-based methods rely on Fourier, Gabor or wavelet transforms to decompose and characterize an image in terms of its frequency or wavelet components (Mallat, 1989). The attributes of barley kernels are derived from the Gabor transform (Feichtinger, 1998; Gabor, 1946), and they describe the magnitudes of frequency components.

In this study, 13 statistical metrics based on the original image and 10 metrics based on image gradient maps were used to describe gray-level distributions. They included the mean, variance, skewness, kurtosis and several percentiles. Eleven grey-level co-occurrence matrix attributes were applied: angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance and difference entropy. All attributes were computed for four directions and three distances between pixels. Five attributes were based on the run-length matrix: short run emphasis, long run emphasis, grey-level nonuniformity, run-length nonuniformity and fraction, and they were computed in four different directions. Moreover, four-parameter autoregressive model and magnitudes of Gabor transform frequency components for 4 directions and 5 scales were used.

Color features are computed from the components of various color models. In this study, the components of the following color models were investigated: RGB (the natural representation for image scanners), YUV, YIQ, HSB, CIE XYZ and CIE Lab. The components of every model carry complementary information. For instance, every component of the HSB model describes different aspects of an image color, such as hue, saturation or brightness. Attributes were statistical features extracted from individual color component histograms (Sangwine and Horne, 1998; Szczypiński et al., 2014) including mean and variance. A total of 20 color attributes were used.

As a result of feature computation, every kernel belonging to data sets S, V and D was characterized by a vector of 295 attributes. The attributes of images from data sets $V+$ and $D+$ were computed within a kernel region and within the wrinkled subregion to produce a total of 590 attributes. The features were computed in MaZda software (Szczypinski et al., 2009).

2.3. Selection of attributes and classification

Computation of a large number of attributes is time consuming, and several-hundred-dimensional spaces are difficult to study. In many cases, only a limited number of features carry information that is relevant for discrimination. What is more, the use of a limited number of features in the classification algorithm is profitable since it reduces the required computational effort.

Therefore, in order to identify these most discriminative attributes, the classifier training stage is preceded by the attributes selection (Blum and Langley, 1997; Kohavi and John, 1997).

There are two general approaches to attribute selection in data mining. In the filter strategy, attribute saliency is estimated individually, and the main advantage is low computational complexity. However, the discriminative power of a feature may be revealed only when it is accompanied by other attributes. Thus, a subset of individually optimized features may prove to be inferior to the actual best solution. In the alternative wrapper approach, the discriminative power of attribute subsets is evaluated. The efficiency of the process depends on the strategy of constructing subsequent feature subsets and the computational complexity of the evaluation routine. It should be noted that testing all possible attribute combinations is not feasible for high-dimensional feature subspaces and computationally demanding evaluation routines.

The proposed feature selection algorithm combines the above mentioned two strategies. The selection begins with estimation of discriminative power of individual features. Subsequently, all combinations of attribute pairs and triples are examined. The tributes are then ordered according to their discrimination ability. For higher-dimensional feature subspaces examination of all possible combinations is extremely time consuming. To overcome this obstacle, the search is limited in time, and the attributes that had been highly ranked so far are examined first.

The discriminative power of an attribute subset is related to the performance of a given classification algorithm. Hence, a set of features selected for a particular classifier may be unsuitable for another classifier. In this study, attributes were selected alternatively for two criteria: Fisher's linear discriminant and balanced accuracy of a nonlinear classifier with hyperellipsoidal decision boundary (HDB).

The first criterion is computed with the use of linear discriminant analysis (LDA) (Fukunaga, 1990) to determines a linear combination of features that best separate vectors of different classes. Fisher's linear discriminant is an optimization criterion defined as the ratio of variance of all vector combinations to the sum of the combination variances in separate classes. If feature vectors inside classes are normally distributed, the best linear discrimination is achieved by maximizing the criterion.

There exist vector distributions that cannot be effectively discriminated by linear classifiers, and the LDA will not yield satisfactory results. Among them, there are distributions in which vectors of one variety (class) form relatively small clusters surrounded by clusters of vectors representing other varieties. Since such clusters cannot be discriminated by a hyperplane, the attribute subspace would be dismissed by LDA as irrelevant. Regardless of the above, the subspace is useful in nonlinear classification where a hypersurface rather than a hyperplane is used as the decision boundary.

A nonlinear HDB classification method for discriminating barley varieties was proposed based on the above findings. The algorithm finds the decision boundary and estimates the discrimination ability of a given feature subspace. The working principle of the algorithm is presented in Figure 3. A principal component analysis (PCA) (Fukunaga, 1990) is performed on vectors of a class to be separated from other classes. The PCA transforms the feature space (or a subspace) into a space of principal components, which are uncorrelated and ordered accordingly to their standard deviations. The decision boundary is a hyperellipsoid with its center located in the center of the cluster, and radii aligned with the principal components and proportional to their standard deviations. The size of hyperellipsoid is adjusted to increase the balanced accuracy of classification and to expand the margin between the vectors of separated classes.

Balanced accuracy estimates the performance of binary classification. It is the average of two ratios: the number of vectors correctly assigned to the first class to the number of vectors actually belonging to the first class, and the number of vectors correctly assigned to the second class to the number of vectors actually belonging to the second class. In the proposed algorithm, balanced

accuracy is a step function of the hyperellipsoid size. The size of the hyperellipsoid is therefore established by determining the center of the widest interval at which balanced accuracy is maximal.

In LDA-driven feature selection, the classes are analyzed in pairs, and the goal of the analysis is to differentiate the two classes. This is a reasonable approach because the vectors of two different classes should form two clusters separated by a hyperplane. Eleven varieties were analyzed in this study, therefore, 55 such pairs had to be investigated which in turn resulted in 55 linear classifiers. Every class is determined by ten of the classifiers. This classifier ensemble was used to identify barley varieties. Every feature vector was presented to all the classifiers. It was identified as a member of a specific class if all the ten classifiers corresponding with the class confirmed its membership.

Unlike the linear classifier, the classifier with a hyperellipsoidal decision boundary should discriminate a vector belonging to a given class from all other vectors. Therefore, the relevant strategy for feature selection and data classification differs from the previously described procedure. In this approach, eleven classifiers are designed, each responsible for identifying one of the eleven varieties.

An artificial neural network (NN) was used as the classification reference method. This choice was motivated by the fact that neural networks are well suited for complex, non-linear and multi-class problems (Chrzanowski et al., 2008; Materka, 1995; Materka and Mizushina, 1996). An advantage of NNs is that carefully trained networks exhibit high generalization capabilities and ensure good performance in training and testing data sets. Among various NN architectures, the multilayer feedforward network (Arbib, 2003; Witten and Frank, 2005) is a widely recognized tool of choice in machine learning, including cereal assessment.

A three-layer network was used in the experiment. A set of attributes obtained by LDA-driven selection was input into the NN. The output layer of the network was determined by the number of classes, and it comprised 11 nodes, each corresponding to one of the eleven classes. The hidden layer was composed of 20 neurons to minimize the risk of overfitting and guarantee high classification accuracy. The scaled conjugate gradient (SCG) method (Møller, 1993) was applied for training. It is a fast algorithm designed specifically for large-scale problems. SCG outperforms the standard back propagation method and many other adaptive optimization techniques, such as conjugate gradient with line-search (CGL) or quasi-Newton learners. Moreover, it does not require any user-dependent parameters and supports fully automated determination of network weights.

All three approaches to varietal classification are supervised learning methods. Those methods generate rules for discrimination, but the suitability of those rules for classifying similar data that have not yet been included in training data sets remains uncertain. Therefore, the utility of the classification scheme have to be validated. For this purpose, a 4-fold cross-validation technique (Witten and Frank, 2005) was applied, and vector sets were randomly divided into four possibly equinumerous subsets. One of the subsets was retained for validation purposes, and the remaining three subsets were used for training. The cross-validation process was repeated four times for each of the four subsets. The results were presented as average classification performance metrics for all four validation tests.

3. Results

The objective of this study was to determine: (i) whether color, texture and shape attributes support image-based discrimination of barley varieties, (ii) whether and to what extent images should be preprocessed to obtain the best classification results, (iii) which feature space reduction technique produces the most discriminating attributes for the classification of barley kernels, and (iv) what is the optimal number of selected features.

To address objectives (ii) and (iii), *S*, *V, D, V*+ and *D*+ data sets were input into the described selection and classification algorithms. The highest values of classification performance indicate the

most effective algorithm and the required level of image preprocessing. Linear classifiers were trained to discriminate each variety against every other variety. A total of 55 pairs were investigated in 5 data sets and 4 folds. Hyperellipsoidal-decision-boundary classifiers were trained to discriminate individual varieties from all the others. Eleven classifiers were trained to identify varieties in 5 data sets and 4 folds. The above resulted in 1320 inferencing processes where the selected attributes and classification rules varied in each case. Moreover, the neural network was also trained on 5 data sets and 4 folds. The resulting data set is enormous, and it cannot be presented due to space constraints, therefore, only representative samples averaged across the folds are shown.

The averaged balanced classification accuracy of the three discussed classifiers obtained from 5 data sets is presented in Table 1. Classification performance increased in successive image preprocessing steps. The highest balanced accuracy of 0.91 was achieved by linear classifier ensembles in the *D*+ data set, and it was higher by 0.04 in comparison with that reported in the *S* data set. The above implies that image preprocessing, which involves kernel side detection, correction of germ-brush orientation and wrinkled area identification, decreases classification error by 4%. A similar tendency was noted in the two remaining classifier techniques – the higher the image preprocessing, the greater the classification accuracy.

To demonstrate that image processing affects classification accuracy, an alternative hypothesis postulating that image processing has no significant impact on classifier performance was formulated. In this case, classification accuracy should randomly increase or decrease as the classifier is first validated on the *S* set and than on the *V*+ or *D*+ set. In the group of 55 linear classifiers, classification performance improved in 40 cases and somewhat deteriorated in 15 cases after the correction of kernel orientation and supplementary analysis of wrinkled regions. The probability that classification performance improved in 40 or more cases (*p*-value) equals 0.00051, which is much below the significance level of 0.01. The above observation disproves the alternative hypothesis and indicates that adjustment of kernel orientation and supplemental analysis of wrinkled regions have significant implications for classification performance.

Averaged classification accuracy of linear and HDB classifiers is presented in Table 2. Balanced accuracy was computed for different dimensionalities of selected feature subspaces to indicate subspace dimensionality with the highest discriminatory power (objective iv). The performance of linear classifier ensembles improved rapidly up to 6-dimensional subspaces. For higher-dimensional subspaces, the observed improvement was less significant, and the highest value of 0.924 was noted for feature space projection. In HDB classifiers, the optimal number of selected attributes is 7 with balanced accuracy of 0.757. A further increase in dimensionality reduces the accuracy.

The confusion matrices of the linear classifier ensemble and the artificial neural network are presented in Tables 3 and 4, respectively. In both cases, columns represent prediction (classification) results and rows represent a true category of kernels. The values express the percentage of kernels of a given variety assigned to the same or another variety. Both linear and NN classifiers were trained on the same selected attribute sets. The applied sets comprised six attributes each, which were selected to maximize Fisher's linear discriminant. The choice of the selection algorithm and the number of attributes in the examined subspaces were justified by the results presented in Table 2.

In linear classifiers, Prymus was the most effectively classified variety where 65% of kernels were correctly identified. Less than 40% of Kormoran kernels were recognized, which was the worst case scenario. The above scores may seem inconclusive for the determination of kernel variety. In practice, however, the final decision is made based on a grain sample consisting of several hundred kernels, by finding a class that is assigned the largest number of elements. The confusion index, defined as the ratio of the highest number of incorrect detections of a single class to the number of correct detections, can be computed to evaluate classification performance. In the linear classifier ensemble, the highest value of the confusion index at 0.27 was reported in the

Serwal class where 12.87% of kernels were incorrectly assigned to the STH variety and where 46.89% of kernels were correctly classified.

The NN classifier is superior to the linear classifier ensemble. It correctly recognized 86% of Blask kernels in the best case scenario and 67% of Kormoran kernels in the worst case scenario. The highest confusion index of 0.10 was observed in the Victoriana class where 7.16% of kernels were incorrectly assigned to the Mercanda variety and more than 70% of kernels were classified correctly.

4. Discussion and Conclusions

The results of this study indicate that determinations of an individual kernel's variety are inconclusive. The error of classification (NN classifier) can be as high as 30%. Yet in practice, there is no need to establish the varietal membership of individual kernels. The variety is usually identified based on a grain sample comprising at least several hundred kernels. Therefore, variety can be determined in relation to the class assigned the highest number of vectors. Consequently, varieties are unlikely to be confused because the number of correct classifications is at least 10 times higher (Table 3 and 4) than the number of incorrect matches. Therefore, it can be concluded that the proposed method based on color, texture and shape attributes is useful for discriminating barley varieties.

Feature selection experiments validated the effectiveness of both selection algorithms. The optimal number of attributes in subsets was determined at seven for the hyperellipsoidal-decision-boundary classifier. The above supports varietal discrimination where the average percentage of correctly identified kernels exceeds 75%. The results produced by the NN classifier are of comparable quality, where 67% - 86% of kernels were correctly identified. The linear classifier ensemble enabled the correct classification of 40% - 65% kernels, which is not impressive. However, individual classifiers can differentiate between two varieties with an average of 91% of correctly identified kernels. This should be considered a very good result.

The optimal number of attributes in subsets was established at 6 or 7. A higher number of attributes did not significantly improve the performance of linear classifiers and deteriorated the performance of the HDB classifiers. No specific attributes or groups of attributes capable of discriminating all varieties could be identified. The subspaces of selected, most discriminating attributes varied between classifiers and training sets. Therefore, we were unable to identify attributes that would be generally considered as most suitable for the varietal classification of barley.

The study demonstrated that image preprocessing significantly influences classification results. The most accurate classification results were obtained in kernel images standardized in terms of anteroposterior and dorsoventral orientation. A supplementary analysis of the wrinkled region also improved the accuracy of successive classification procedures.

The results of the study are promising, and they should pave the way to further research with the aim of developing an effective model for successive growing seasons. A detailed methodology for preparing grain for analysis should be also developed. It should enhance the quality of scanned kernel images and expose surface texture attributes specific to a given variety. Grain could also be sorted before scanning to eliminate undeveloped kernels and objects of inadequate shape, color or texture that could affect classification quality.

Acknowledgments

The authors wish to thank the Polish Ministry of Science and Higher Education for financial support as part of grant No. 4498/B/P01/2010/39. The following computer programs were used in the study: Ziarna (https://gitorious.org/ziarna) for kernel image preprocessing and MaZda (https://gitorious.org/qmazda) for attribute computation and feature selection.

References

Arbib, M.A., 2003. The Handbook of Brain Theory and Neural Networks. MIT Press.

Berthold, M., Hand, D.J. (Eds.), 2003. Intelligent data analysis: an introduction, 2nd rev. and extended ed. ed. Springer, Berlin ; New York.

Blum, A.L., Langley, P., 1997. Selection of relevant features and examples in machine learning. Artif. Intell. 97, 245–271.

Chrzanowski, L., Drozdz, J., Strzelecki, M., Krzeminska-Pakula, M., Jedrzejewski, K.S., Kasprzak, J.D., 2008. Application of Neural Networks for the Analysis of Intravascular Ultrasound and Histological Aortic Wall Appearance-An In Vitro Tissue Characterization Study. Ultrasound Med. Biol. 34, 103–113. doi:10.1016/j.ultrasmedbio.2007.06.021

Feichtinger, H.G., 1998. Gabor analysis and algorithms: Theory and applications. Birkhauser Boston.

Fukunaga, K., 1990. Introduction to statistical pattern recognition. Academic Press.

Gabor, D., 1946. Theory of communications. J. Inst. Electr. Eng. 93, 429–457.

Goyal, S., 2013. Predicting properties of cereals using artificial neural networks: A review. Sci. J. Crop Sci. Vol 2 No 7 2013 July.

Haralick, R.M., 1979. Statistical and structural approaches to texture. Proc. IEEE 67, 786–804.

Haralick, R.M., Shanmugam, K., Dinstein, I., 1973. Textural Features for Image Classification. Syst. Man Cybern. IEEE Trans. On 3, 610–621.

Hulasare, R., Dronzek, B., Jayas, D., 2003. Grain-Grading Systems, in: Ramaswamy, H., Vijaya Raghavan, G., Chakraverty, A., Mujumdar, A. (Eds.), Handbook of Postharvest Technology. CRC Press, pp. 41–55.

Jain, A.K., 1989. Fundamentals of digital image processing. Prentice-Hall, Inc.

Jain, A.K., Duin, R.P.W., Jianchang Mao, 2000. Statistical pattern recognition: a review. Pattern Anal. Mach. Intell. IEEE Trans. On 22, 4–37.

Kohavi, R., John, G.H., 1997. Wrappers for feature subset selection. Artif. Intell. 97, 273–324. doi:doi: 10.1016/S0004-3702(97)00043-X

Majumdar, S., Jayas, D.S., 2000. Classification of cereal grains using machine vision: I. Morphology models. Trans. ASAE 43, 1669–1675.

Mallat, S.G., 1989. Multifrequency channel decompositions of images and wavelet models. Acoust. Speech Signal Process. IEEE Trans. On 37, 2091–2110.

Manickavasagan, A., Sathya, G., Jayas, D.S., White, N.D.G., 2008. Wheat class identification using monochrome images. J. Cereal Sci. 47, 518–527. doi:10.1016/j.jcs.2007.06.008

Materka, A., 1995. Neural network technique for parametric testing of mixed-signal circuits. Electron. Lett. 31, 183–184. doi:10.1049/el:19950148

Materka, A., Mizushina, S., 1996. Parametric signal restoration using artificial neural networks. IEEE Trans. Biomed. Eng. 43, 357–372. doi:10.1109/10.486256

Materka, A., Strzelecki, M., 1998. Texture analysis methods-a review (COST B11). Technical University of Lodz, Institute of Electronics.

Mebatsion, H.K., Paliwal, J., Jayas, D.S., 2013. Automatic classification of non-touching cereal grains in digital images using limited morphological and color features. Comput. Electron.

Agric. 90, 99–105. doi:10.1016/j.compag.2012.09.007
Møller, M.F., 1993. A scaled conjugate gradient algorithm for fast supervised learning. Neural Netw. 6, 525–533. doi:http://dx.doi.org/10.1016/S0893-6080(05)80056-5
Neuman, M.R., Sapirstein, H.D., Shwedyk, E., Bushuk, W., 1989. Wheat grain colour analysis by digital image processing I. Methodology. J. Cereal Sci. 10, 175–182. doi:10.1016/S0733-5210(89)80046-3
Neuman, M., Sapirstein, H.D., Shwedyk, E., Bushuk, W., 1987. Discrimination of wheat class and variety by digital image analysis of whole grain samples. J. Cereal Sci. 6, 125–132. doi:10.1016/S0733-5210(87)80049-8
Paliwal, J., Shashidhar, N.S., Jayas, D.S., 1999. Grain kernel identification using kernel signature. Trans. ASAE 42, 1921–1924.
Pavlidis, T., 1980. Algorithms for shape analysis of contours and waveforms. IEEE Trans. Pattern Anal. Mach. Intell. 2, 301–312.
Sangwine, S.J., Horne, R.E.., 1998. The colour image processing handbook. Chapman & Hall.
Szczypiński, P., Klepaczko, A., Pazurek, M., Daniel, P., 2014. Texture and color based image segmentation and pathology detection in capsule endoscopy videos. Comput. Methods Programs Biomed. 113, 396–411.
Szczypinski, P.M., Strzelecki, M., Materka, A., Klepaczko, A., 2009. MaZda - A software package for image texture analysis. Comput. Methods Programs Biomed. 94, 66–76. doi:doi: 10.1016/j.cmpb.2008.08.005
Szczypiński, P.M., Zapotoczny, P., 2012. Computer vision algorithm for barley kernel identification, orientation estimation and surface structure assessment. Comput. Electron. Agric. 87, 32–38.
Visen, N.S., Paliwal, J., Jayas, D.S., White, N.D.G., 2002. AE—Automation and Emerging Technologies: Specialist Neural Networks for Cereal Grain Classification. Biosyst. Eng. 82, 151–159.
Witten, I.H., Frank, E., 2005. Data Mining. Practical Machine Learning Tools and Techniques. Morgan Kaufmann.
Zapotoczny, P., 2011a. Discrimination of wheat grain varieties using image analysis and neural networks. Part I. Single kernel texture. J. Cereal Sci. 54, 60–68.
Zapotoczny, P., 2011b. Discrimination of wheat grain varieties using image analysis: morphological features. Eur. Food Res. Technol. 233, 769–779.
Zielinska, M., Zapotoczny, P., Białobrzewski, I., Zuk-Golaszewska, K., Markowski, M., 2012. Engineering properties of red clover (Trifolium pratense L.) seeds. Ind. Crops Prod. 37, 69–75. doi:10.1016/j.indcrop.2011.12.002

Table 1. Average balanced classification accuracy for various steps of image processing. Six-dimensional subspaces were used.

Classifier	*S*	*V*	*D*	*V*+	*D*+
HDB	0.732	0.729	0.743	0.732	**0.755**
LDA	0.868	0.871	0.896	0.873	**0.910**
NN	0.869	0.875	0.884	**0.895**	0.893

The most adequate results are indicated with bold.

Table 2. Average balanced classification accuracy versus dimensionality of selected subspace. The *D*+ data set was used.

Classifier	1D	2D	3D	4D	5D	6D	7D	8D	9D	10D	591
LDA	0.782	0.852	0.876	0.891	0.903	**0.910**	0.912	0.914	0.915	0.917	0.924
HDB	0.636	0.703	0.732	0.747	0.754	**0.755**	**0.757**	0.754	0.755	0.75	0.506

The most adequate results are indicated with bold.

Table 3. Confusion matrix of the linear classifiers' ensemble. The *D*+ data set and six-dimensional subspaces were used.

		Classification result										
	Variety	Afrodita	Blask	Bordo	Conchita	Kormoran	Mercanda	Prymus	Serwal	Signora	STH	Victoriana
True category	Afrodita	**54.99**	0.15	0.30	4.73	3.70	0.15	2.22	5.77	0.74	6.07	0.74
	Blask	1.04	**63.17**	2.37	1.78	1.18	2.37	3.11	0.74	2.07	1.78	0.74
	Bordo	3.70	3.85	**54.59**	0.30	2.22	2.22	5.03	0.30	1.33	1.48	2.96
	Conchita	5.18	2.07	0.44	**61.54**	4.88	0.00	0.30	7.25	0.89	1.33	0.00
	Kormoran	3.85	2.51	1.92	5.33	**39.94**	3.11	7.84	6.66	0.59	2.37	1.33
	Mercanda	0.89	6.07	2.81	0.00	2.51	**55.47**	5.62	0.00	3.85	1.92	2.66
	Prymus	0.15	0.74	2.37	0.00	3.55	1.63	**65.09**	0.30	2.96	1.18	1.78
	Serwal	6.21	0.89	1.18	5.92	3.99	0.00	0.59	**46.89**	0.44	**12.87**	0.15
	Signora	1.63	2.51	2.81	1.63	0.44	2.96	2.51	1.92	**47.93**	7.69	5.33
	STH	6.21	0.30	1.48	1.92	3.55	1.04	3.70	8.73	2.22	**43.20**	1.18
	Victoriana	2.51	1.92	4.29	1.04	1.63	4.44	5.77	0.59	6.95	2.22	**44.82**

Correct classification rates and the highest incorrect classification rate are indicated with bold.

Table 4. Confusion matrix of the artificial neural network. The *D*+ data set and six-dimensional subspaces were used.

		Classification result										
	Variety	Afrodita	Blask	Bordo	Conchita	Kormoran	Mercanda	Prymus	Serwal	Signora	STH	Victoriana
True category	Afrodita	**76.46**	0.00	0.50	6.60	3.86	0.87	0.75	5.35	1.25	3.61	0.75
	Blask	0.00	**86.10**	2.86	1.09	1.36	2.45	1.09	1.09	1.23	1.09	1.63
	Bordo	0.80	1.61	**73.95**	0.64	2.89	2.73	2.41	1.29	6.59	1.61	5.47
	Conchita	6.65	0.81	0.14	**76.53**	6.51	0.27	0.54	5.83	0.14	2.04	0.54
	Kormoran	4.22	1.74	2.61	5.71	**67.33**	1.61	2.98	4.97	2.86	3.98	1.99
	Mercanda	0.54	3.79	3.79	0.81	1.22	**74.83**	4.47	0.14	3.11	1.76	5.55
	Prymus	0.81	0.46	2.21	0.70	2.90	3.14	**79.21**	0.70	3.48	1.74	4.65
	Serwal	2.81	1.34	0.80	8.16	5.21	0.13	0.00	**73.66**	1.07	6.42	0.40
	Signora	2.27	1.81	6.50	0.45	3.63	3.47	5.29	0.60	**69.18**	3.63	3.17
	STH	4.01	1.60	1.49	2.06	2.98	1.83	1.49	6.76	2.41	**73.77**	1.60
	Victoriana	1.49	1.35	5.41	0.41	2.43	**7.16**	6.49	0.27	3.65	1.22	**70.14**

Correct classification rates and the highest incorrect classification rate are indicated with bold.

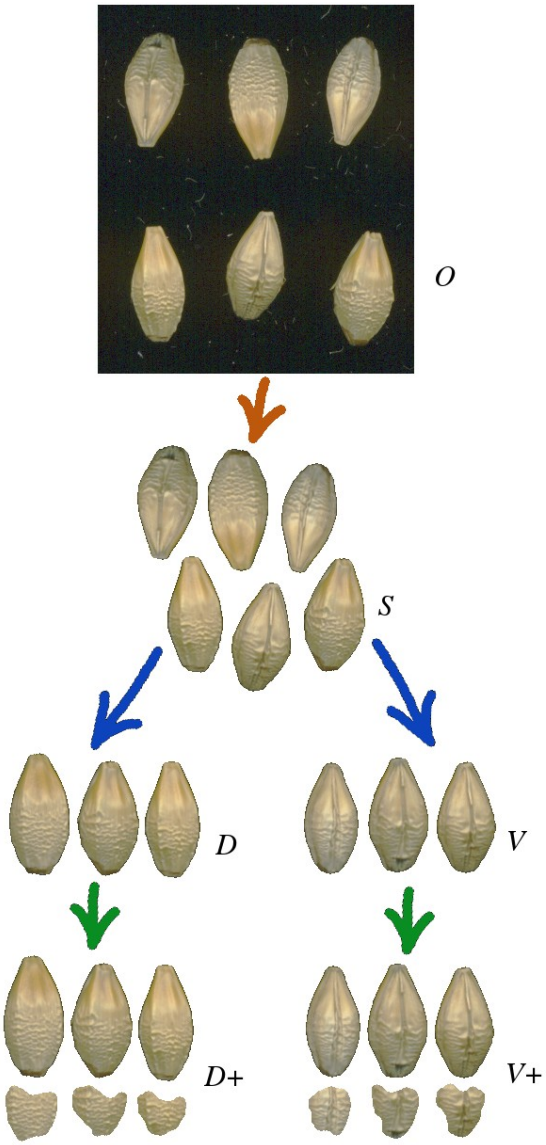


Fig. 1. Image preprocessing diagram and data sets for analysis: *O* – fragment of the original image, *S* – kernels identified after image segmentation, *D* – adjusted orientation of the dorsal view, *V* – adjusted orientation of the ventral view, *D*+ – dorsal view and the corresponding wrinkled areas, *V*+ – ventral view and the corresponding wrinkled areas.

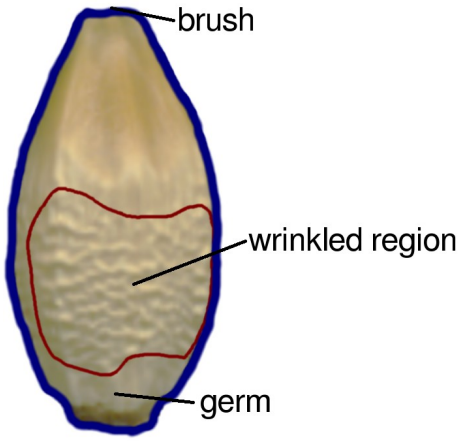
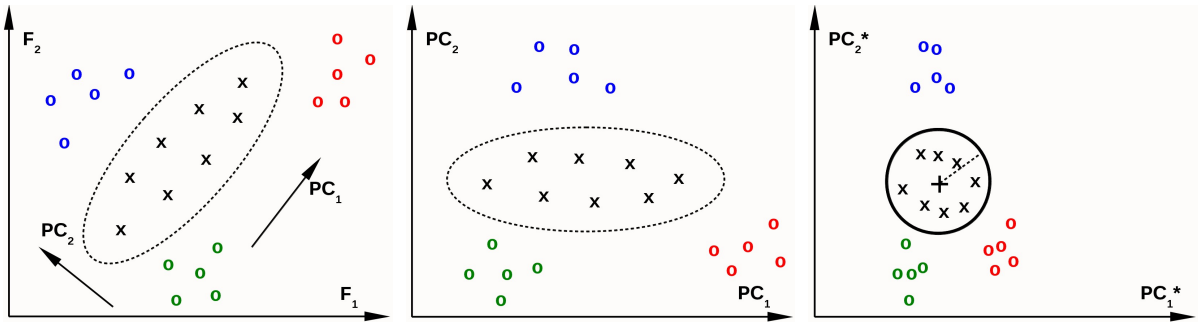
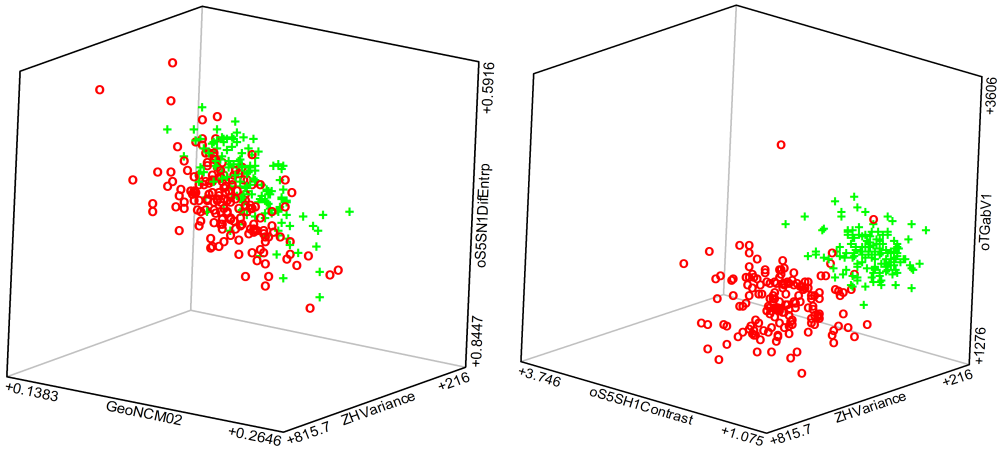


Fig. 2. Kernel image with an indication of the estimated location of the germ, brush and the wrinkled region.



(a) (b) (c)

Fig. 3. The hyperellipsoidal-decision-boundary classification concept in two dimensional space: (a) space of original features F1 and F2, (b) space of two principal components, PC1 and PC2, obtained by PCA and (c) space of principal components after normalization of variances. Vectors of the discriminated class are labeled with 'x', vectors of other classes are labeled with 'o'; '+' indicates the center of a separated vector cluster.



(a) (b)

Fig. 4. Vector distributions in (a) Serwal-STH and (b) Conchita-Prymus classes in the most discriminating 3D subspaces of data sets *D*+. Serwal and Conchita vectors are labeled with 'o' and STH and Prymus vectors are labeled with '+'.